

深層学習に物理モデルを結合させた排水機場水位予測手法の構築
 Developing a water level prediction method that combines deep learning with a physical model
 at a drainage pumping station.

○木村延明*・皆川裕樹*・福重雄大*・馬場大地**

○Nobuaki KIMURA・Hiroki MINAGAWA・Yudai FUKUSHIGE・Daichi BABA

1. はじめに

近年、リアルタイム予測などに導入されるデータ駆動型の深層学習モデルは、正確かつ大量のデータを利用することで、精度の高い予測モデルとなる。本研究対象である排水機場調整池の水位予測に深層学習モデルを適用させる場合、観測データの異常値除去に手間がかかり、また長期にわたる大量データの収集も容易ではない。たとえ正確で大量なデータが準備できたとしても、過去データの中から学習できない未経験の事象は、予測そのものができない。これを改善するために、物理モデルを導入する。物理モデルから仮想的に生成される大量の模擬データ（未経験の事象も含む）を準備し、それを学習することで不十分な観測データを補うことができる。しかし、一般的に物理モデルは近似解を提供するものなので、観測データのもつパターン（特徴量）により近づける必要がある。これを可能にするために、あるデータの特徴量を別のデータに転移させる「転移学習」¹⁾を導入する。本研究の目的は、物理モデルから提供される模擬データを用いて事前学習した深層学習モデルに転移学習を導入することで、未経験の事象も含めてより現実的な予測を可能にする手法を構築することである。

2. 方法

本研究の予測手法は、常時排水を行う低平農地の排水機場調整池の水位予測を対象に、図-1 に示す 3 つのステップで構築される。(1) 現地観測された気象データに基づき、確率モデルから生成された豪雨時を含む 1000 個分の模擬降雨 (300 mm/72 h)²⁾ を用いて、内水氾濫解析モデル(物理モデル)を駆動させ、仮想的な気象状況下における模擬水位データを大量に生成する。この模擬データは、観測水位データの代替データと見なす。なお、300 mm/72 h は、対象にした期間最大の降雨量に匹敵する降雨イベントである。(2) 物理モデルから得られた模擬水位データを用いて、深層学習モデルの事前学習を行う。(3) 事前学習された深層学習モデルを用いて、模擬水位デー

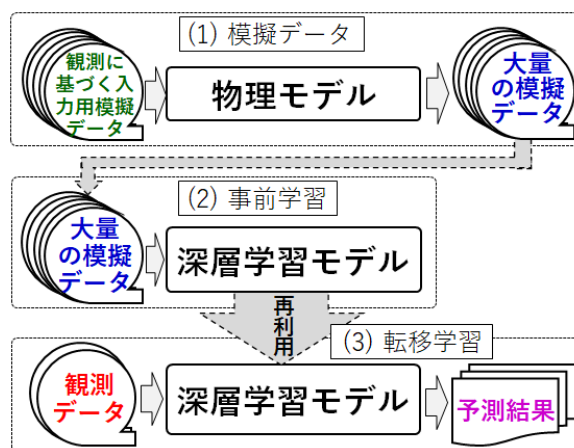


図-1 提案する予測手法の概念とデータフロー
 Concept & data flow of our proposed methodology.

* 農研機構－農村工学研究部門 The Institute for Rural Engineering, NARO

** (株)アーク情報システム ARK Information Systems, INC.

キーワード 排水機場水位予測, 深層学習モデル, 物理モデル, 転移学習

タの特徴量を観測データに取り込むために転移学習を導入する。その上で、水位予測を行い、結果の精度検証を実施する。なお、転移学習では、事前学習された深層学習モデルの内部パラメータを一部分再学習して最適化する。観測データは、約7年間半に測定された水位データで、常時排水時のデータが98%以上を占めるものを利用する。300 mm/72 hの降雨量に対応する模擬水位データは、降雨期間の72時間に水位波形のテールを計算するために2日間を加えた120時間分のシミュレーション結果である。10分割交差検証法を用いて予測精度の評価を行い、二乗平均平方根誤差 (RMSE) と Nash-Sutcliffe 係数 (NS) で定量評価する。深層学習モデルは、時間予測に有用な長短期記憶 (LSTM³⁾) を採用した。

3. 予測結果とまとめ

1000個の洪水イベントを模擬した水位データで事前学習させたLSTMを転移学習によって観測データに適用した。再学習回数は、誤差の変化が安定した50回に設定した。期間最大の洪水イベントが含まれる分割区間でのリードタイム1時間の予測結果を観測データと共に図-2に示す。また、転移学習なしの(模擬水位データで事前学習しない)LSTMの予測結果も同時に示す。転移学習あり・なしの誤差の定量評価は、この常時排水を含む分割区間において共にRMSE=0.025, NS=0.85で、ほぼ同程度の精度を示した。期間最大の洪水イベントの水位予測(図-2の右の小パネル)では、転移学習ありの予測結果は、明らかにピークをより良く再現できた。転移学習あり/なしの定量評価では、それぞれRMSE=0.07/0.08 m, NS=0.95/0.94と若干ではあるが誤差の改善が見られた。以上から、物理モデルから得られた多くの洪水イベントの模擬水位データを事前に学習したLSTMについて、転移学習を導入し、常時排水を含む観測水位データに適用することで、期間最大の洪水イベントの良好な再現に加え、常時排水を含む区間でも転移学習なしの予測結果と同程度の予測精度が得られた。

謝辞: 本研究は環境研究総合推進費(JPMEERF20S11803)により実施させて頂いた。

参考文献: 1) Pratt (1992) In Proc. of the Advances in NIPS 5, pp. 154-196; 2) 皆川ら (2020) 土木学会論文集B1(水工学), 76(2), pp. I_349-I_354; 3) Hochreiter and Schmidhuber (1997) Long short-term memory. Neural Comput., 9(8), pp. 1735-1780.

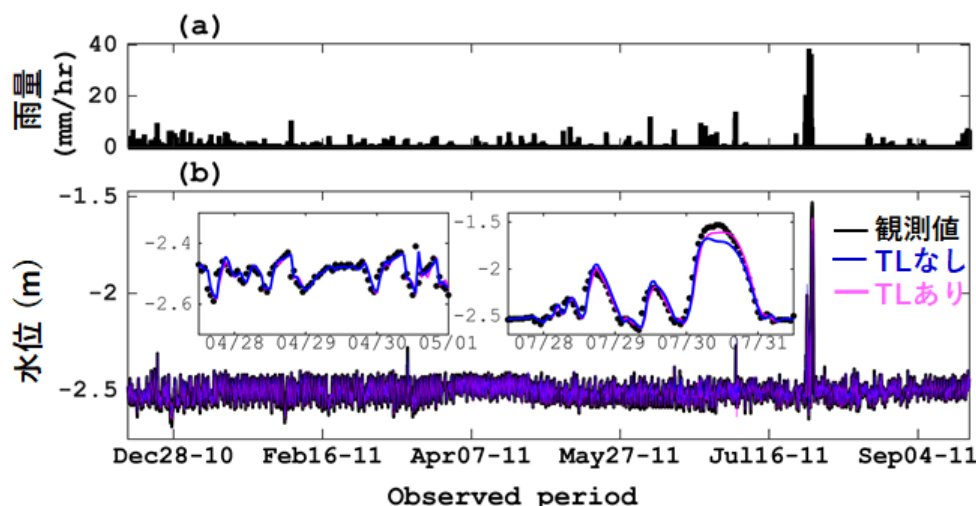


図-2 期間最大の水位高を含む分割区間のリードタイム1時間の予測結果(小パネル:左-常時と右-洪水時の拡大図)と雨量, TL=Transfer learning (転移学習)。Predicted water levels in 1 h lead time with/without TL, compared with observation and rainfall in the segmented period, including the maximum flood event. Small panels show the water levels during a typical drain operation (left) and a flood event.